PCI-SIG ENGINEERING CHANGE NOTICE

TITLE:	PME_Turn_Off Protocol
DATE:	13 Feb 2004 (updated 11 Mar 2004)
AFFECTED DOCUMENT:	PCI Express Base Specification Revision 1.0a
SPONSOR:	David Harriman, Intel Corporation

Part I

1. Summary of the Functional Changes

This document proposes changes to the PME_Turn_Off protocol to improve consistency, robustness and utility.

Note that this document comprehends preceding errata.

2. Benefits as a Result of the Changes

Improves consistency, robustness and utility of PME_Turn_Off protocol..

3. Assessment of the Impact

Existing endpoints may require modification to be compliant with these changes.

4. Analysis of the Hardware Implications

Enables hardware notification of impending power removal regardless of software operations, but potentially increases design and validation complexity to handle new scenarios.

5. Analysis of the Software Implications

PCI Express aware software unaffected. Non-PCIe aware software should operate more smoothly.

Part II

Detailed Description of the change

In Section 5.2 Link State Power Management, edit as shown:

The following example sequence, leading up to entering a system sleep state, illustrates the multi-step Link state transition process:

1. System Software directs all functions of a Downstream component to $D3_{hot}$.
2. The Downstream component then initiates the transition of the Link to L1 as required.
3. System Software then causes the Root Complex to broadcast the PME_Turn_Off Message in preparation for removing the main power source.
4. This Message causes the subject Link to transition back to L0 in order to send it, and to enable the Downstream component to respond with PME_TO_Ack.
5. After the PME_TO_Ack is sent, the Downstream component then initiates the L2/L3 Ready transition protocol.

L0 → L1 → L0 → L2/L3 Ready

As the following example illustrates, it is also possible to remove power without first placing all devices into $D3_{hot}$:

1. System Software causes the Root Complex to broadcast the PME_Turn_Off Message in preparation for removing the main power source.
2. The Downstream components respond with PME_TO_Ack.
3. After the PME_TO_Ack is sent, the Downstream component then initiates the L2/L3 Ready transition protocol.

L0 → L2/L3 Ready

In Section 5.3.1.4.1 $D3_{hot}$ State:

When a function is in $D3_{hot}$, it must respond to configuration accesses targeting it. ~~They must also participate in the PME_Turn_Off/PME_TO_Ack protocol. Refer to Section 5.3.3 details.~~ Once in $D3_{hot}$ the function can later be transitioned into $D3_{cold}$ (by removing power from its host component).

~~Transitions into the $D3_{hot}$ state are used to establish a standard process for graceful saving of functional state immediately prior to entering a deeper power savings state where power is removed.~~

...

In Section 5.3.2. PM Software Control of the Link Power Management State:

The power management state of a Link is determined by the D-state of its Downstream component.

Table 5-2 depicts the relationships between the power state of a component (Endpoint or Switch) and its Upstream Link.

Table 5-2: Relation Between Power Management States of Link and Components

Downstream Component D-State	Permissible Upstream Component D-State	Permissible Interconnect State
D0	D0	L0, L0s, L1 (1), L2/L3 Ready
D1	D0-D1	L1, L2/L3 Ready
D2	D0-D2	L1, L2/L3 Ready
$D3_{hot}$	D0-$D3_{hot}$	L1, L2/L3 Ready~~(2)~~
$D3_{cold}$	D0-$D3_{cold}$	L2 ~~(23)~~, L3

Notes:

1. All PCI Express components are required to support ASPM with L0s entry during idle at a minimum. The use of L1 within D0 is optional.

~~2. When all functions within a Downstream component are programmed to $D3_{hot}$ the Downstream component must request the transition of its Link to the L1 state using the PM_Enter_L1 DLLP. Once in $D3_{hot}$, following the execution of a PME_Turn_Off/PME_TO_Ack handshake sequence, the Downstream component must then request a Link transition to L2/L3 Ready using the PM_Enter_L23 DLLP. Following the L2/L3 Ready entry transition protocol the Downstream component must be ready for loss of main power and reference clock.~~

2. If Vaux is provided by the platform, the Link sleeps in L2. In the absence of Vaux, the L-state is L3

The following rules relate to ~~conditions governing Link state transition in the software directed~~ PCI-PM compatible power management ~~scheme are defined as~~:

- Devices in D0, D1, D2, and D3hot must respond to the receipt of a PME_Turn_Off Message by the transmission of a PME_TO_Ack Message.
- In any device D state, following the execution of a PME_Turn_Off/PME_TO_Ack handshake sequence, a Downstream component must request a Link transition to L2/L3 Ready using the PM_Enter_L23 DLLP. Following the L2/L3 Ready entry transition protocol the Downstream component must be ready for loss of main power and reference clock.
- A Switch or single function Endpoint device must initiate a Link state transition of its Upstream Port (Switch) or Port (Endpoint), to L1 based solely upon that Port being programmed to D1, D2, or $D3_{hot}$. In the case of the Switch, system software bears the responsibility of ensuring that any D-state programming of a Switch's Upstream Port is done in a compliant manner with respect to PCI Express hierarchy-wide PM policies (i.e., the Upstream Port cannot be programmed to a D-state that is any less active than the most active Downstream Port and Downstream connected component/function(s)).

- Multi-function Endpoints must not initiate a Link state transition to L1 until all of their functions have been programmed to a non-D0 D-state.

In Section 5.3.2.3. Entry into the L2/L3 Ready State:

....

- ~~The Downstream component must be in $D3_{hot}$ prior to being transitioned into the L2/L3 Ready state, i.e., a PME_Turn_Off Message must never be sent unless all functions Downstream of its point of origin are currently in $D3_{hot}$.~~

  ~~In contrast, a Downstream component initiating a transition to L1 would have always initially been in D0, and had just been reprogrammed to D1, D2, or $D3_{hot}$.~~

In Section 5.3.3.2.1 PME Synchronization (Note: This text reflects changes in Errata C16):

...

All Endpoints must accept and acknowledge the PME_Turn_Off Message regardless of the D state of the associated device or any of its functions for a multi function device ~~from within the D3hot State~~. Once an Endpoint has sent a PME_TO_Ack Message, it must then prepare for removal of its power and reference clocks by initiating a transition to the L2/L3 Ready state.

...

In Section 5.3.3.4. PME Rules:

...

- On receiving a PME_Turn_Off Message ~~while in the $D3_{hot}$ state~~, the device must block the transmission of PM_PME Messages and transmit a PME_TO_Ack Message upstream. The component is permitted to send a PM_PME Message after any one of the following has occurred: the entrance and exit of L2/L3 Ready, the receipt of any TLP, fundamental reset (including the case of removal and restoration of main power to the component).

-- end of change --